

Attending to different levels of structure in a visual image

R. A. KINCHLA and VICTOR SOLIS-MACIAS
Princeton University, Princeton, New Jersey

and

JAMES HOFFMAN
University of Delaware, Newark, Delaware

Two experiments are reported in which observers had to utilize information from one of two structural levels of visual stimulus patterns (large letters composed of smaller ones). They could utilize information more rapidly from one level only at the cost of slower utilization from the other. This tradeoff defines an empirical attention operating characteristic (AOC) which is consistent with a simple mathematical model of the perceptual process: when viewing a stimulus, the observer selects one of two alternative "attentional" strategies, where each strategy is optimal for utilizing information from one structural level, but less than optimal for the other.

Most people feel that they can attend to specific levels of structure in a visual image: for example, choose to "see" either a whole ("higher level") form, such as a "face," or some constituent ("lower level") form, such as an "eye" or a "nose." Of course, one's pattern of eye movements is undoubtedly different when one tries to see a "face" rather than some constituent of that face. Yet, even if an image is presented so briefly (e.g., 100 msec) that only one fixation is possible, the phenomenal impression of being able to attend to one level of form or another persists. This suggests that perception of the different structural levels of an image may involve covert attentional (selective) processes as well as overt eye movements. This paper presents experimental evidence and a theoretical model consistent with this view.

Image Structure and Structural Redundancy

Modern theories of visual perception have been strongly influenced by work on computer pattern recognition, particularly the concept of image structure employed in syntactic scene analysis (see Fu, 1974). The idea is that an image ("scene") may be parsed into hierarchical levels of form, much as a paragraph can be parsed into sentences, phrases, words, etc. For example, a "face" may be decom-

posed into lower order forms such as "eyes," "ears," "nose," etc., which in turn may be decomposed into even lower order forms, such as an "eye" into "pupil," "lids," "lashes," etc. The value of such a representation in computer pattern recognition stems from the normal redundancy (correlation) between one level of structure and another ("faces" are likely to contain "eyes," just as "eyes" imply the presence of a "face"). This structural redundancy can facilitate computer recognition of scenes in the same way that syntactical redundancy can facilitate the processing of language, recognition of a form at one level serving as a clue for subsequent recognition of both higher and lower level forms.

Human visual perception is clearly influenced by structural redundancy. For example, Palmer (1975) showed that one's perception of a tachistoscopically presented form can be strongly influenced by the scene in which it is embedded. In one case, subjects saw the same object as a "loaf of bread," when embedded in a "kitchen" scene, or as a "mailbox," when seen in the context of a "country road." Naturalistic scenes of this sort provide strong cues for perception of constituent forms because people are familiar with the normal structure of such scenes. However, this is also why it is difficult to manipulate the structural redundancy in natural scenes; for example, there is a biologically appropriate relation between a "face" and its parts which cannot be arbitrarily modified. This is why Kinchla (1974) proposed using the sort of experimental stimuli illustrated in Figure 1, large letters composed of smaller ones. Such stimuli have two levels of structure, a global form (the large letter) and local forms (the constituent, smaller letters), both of which are familiar and

This work was supported under NSF Grant BNS 80-04649 held by the senior author, and was reported earlier in a paper presented at the 1979 meeting of the Psychonomic Society in San Antonio, Texas. R. A. Kinchla's mailing address is: Department of Psychology, Green Hall, Princeton University, Princeton, New Jersey 08544.

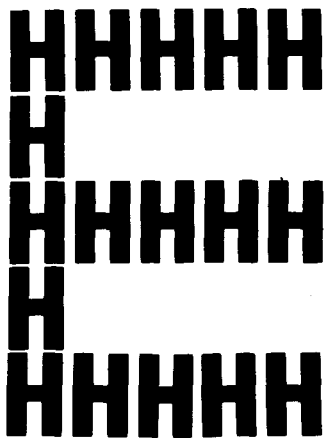


Figure 1. A stimulus pattern of the sort originally proposed by Kinchla (1974) which has at least two levels of structure: a "global form" (H) and many instances of a "local form" (E).

easily recognized. However, in contrast to most naturalistic scenes, the degree of structural redundancy (correlation between levels of form) can be arbitrarily manipulated. For example, Kinchla (1977) varied the probability that specific global forms (large letters) would contain particular local forms (small letters). When viewed tachistoscopically, the redundant information in the global forms systematically influenced perception of the local forms: small target letters were detected more often in large letters likely to contain them than in those not likely to contain them.

Order of Processing Structural Levels

The various components of an image are probably not processed equally fast. Forms that are nearer the fovea, more familiar, simple, distinctive, expected, etc., are probably recognized more rapidly than others. Speed or order of processing may also be related to structural level. There are at least three possibilities for such an ordering: lower levels first followed by progressively higher levels, a "bottom-up" sequence; higher levels first followed by progressively lower levels, a "top-down" sequence; or some intermediate level first followed by processing of both progressively higher and progressively lower levels, a "middle-out" sequence.

Bottom-Up Processing

Many of the early theories of human pattern recognition (e.g., Rumelhart, 1970) incorporated a bottom-up sequence of processing. Lower level "features" (lines, points, edges, etc.) were processed first, then synthesized into progressively higher order forms. This view was encouraged both by similar constructive algorithms used in computer pattern recognition and by electrophysiological studies of

receptive fields which seemed to show simple feature detecting fields at the periphery of the visual system feeding into progressively more complex fields as one moved toward the visual cortex.

Top-Down Processing

Navon (1977), on the other hand, argued that visual processing is not bottom-up, but is *invariably* top-down ("global-to-local"). His principal evidence for this view came from reaction-time experiments involving stimuli similar to the one shown in Figure 1. Subjects were sometimes asked to respond on the basis of the large-letter, "global instructions," and sometimes on the basis of the smaller constituent letters, "local instructions." While they seemed able to ignore the small letters under the global instructions, their response latencies were influenced by the large letter under the local instructions (they were slower when the global form indicated a different response from the local one). Navon argued that this showed that global forms were always processed first: subjects could respond to global forms before the more slowly processed local forms produced interference, while the opposite was impossible.

However, subsequent studies employing similar experimental tasks (Hoffman, 1980; Miller, 1982a) have shown *both* global and local interference effects (neither level could be completely ignored), as if both levels were processed at about the same speed. This is also the conclusion Pomerantz and Sager (1975) drew from their earlier study in which observers simply sorted patterns similar to that in Figure 1. Neither the global nor the local forms could be completely ignored, although (contrary to Navon's results) subjects were better at ignoring the global forms. Finally, Boer and Keuss (1982) employed a deadline procedure to evaluate speed-accuracy trade-offs on a task similar to Navon's. Again, even under extreme speed pressure (lowest accuracy levels) their observers were unable to completely ignore either level of form.

Middle-Out Processing

Kinchla (Kinchla & Wolfe, 1979) has argued that processing order is neither consistently bottom-up nor consistently top-down. Rather, forms at some intermediate level of structure, subtending some optimal region of the visual field, are recognized first, with subsequent recognition of both higher and lower level forms. For example, the first component of an image to be recognized might be a "face," with rapid subsequent recognition of both higher level forms, such as a "whole person," and lower level forms, such as an "eye" or "nose." An experiment supporting this "middle-out" view of processing order was reported by Kinchla and Wolfe (1979). Subjects were shown stimulus patterns similar to that in Figure 1 and asked to quickly decide whether a par-

ticular target letter was defined at either the global or local level. The relative speed with which they detected targets at each level depended on the angular size of the stimulus pattern: when the height of the entire pattern subtended less than about 7 deg visual angle, global targets were detected faster, but when this angle exceeded 7 deg local targets were detected faster. Since Navon (1977) only used stimuli smaller than 7 deg, his finding of an "invariant" global precedence would seem to be an artifact of his restricted angular variation.

Martin (1979) found evidence for both "global" and "local" precedence depending on what she termed the "sparsity" of local letters. Her global letters were a fixed size but defined either as a subset of a 5×7 matrix of small letters or a 3×5 matrix of slightly larger letters, a difference in the "sparsity" of local forms according to Martin. However, her sparsity manipulation is entirely confounded with variation in the relative angular sizes of the global and local forms. Thus, her results are consistent with the effects of angular variation reported earlier by Kinchla and Wolfe.

Hoffman (1980) has shown that the relative speed with which subjects report targets defined at the global or local level can be manipulated by degrading (distorting) the forms at either level. This result again suggests that processing is not consistently "top-down" or "bottom-up," but rather that forms at some intermediate level of optimal visibility (due to angular size, retinal locus, or degree of distortion, etc.) are processed first, with subsequent processing of both higher and lower levels of form. Note that processing of different structural levels may initially proceed simultaneously and independently. However, as soon as enough information is processed from one level, it should begin to influence subsequent processing at other levels because of the redundancy between different levels of structure in most natural scenes. Again, the particular levels of structure processed first will depend on viewing distance, familiarity, etc. A more extensive discussion of this view and of the relation between structural hierarchy and spatial frequency bands is presented in Kinchla and Wolfe (1979).

"Processing" vs. Overt Utilization

Up to this point, we have intentionally used the term "processing" somewhat loosely. This is because such terms can really be clearly defined only in the context of an explicit theoretical model. Yet a variety of models are suggested in the previously cited literature, some explicitly and others only vaguely implicit in the author's verbal characterization of the perceptual process. For the present it seems best to defer further discussion of theoretical mechanisms. We shall simply consider how rapidly an observer can reflect or utilize information about various aspects of

a stimulus in the form of some overt response. We shall refer to this as "*overt information utilization*." Thus, in less theoretical terms, the previously cited experiments by Hoffman (1980), Kinchla and Wolfe (1979), and Martin (1979) show that relative angular size and degree of distortion influence how rapidly an observer can utilize information from different structural levels, and the interference effects found by Boer and Keuss (1982), Hoffman (1980), Martin (1979), Miller (1982a), Navon (1977), and Pomerantz and Sager (1975) show that the speed at which information from one level can be utilized may depend on information at another level even when one is told to ignore it and responding is very rapid.

Volitional Control of Processing Order

Clearly, then, various physical aspects of a stimulus affect how rapidly information from the different structural levels of an image may be utilized. Other factors, such as an observer's "expectancies" or his relative familiarity with various image components, are also important. But suppose such factors were held constant. Does an observer have any volitional control of the structural level he may utilize most rapidly? Selectivity of this sort would constitute a type of *attentional* process.

A general approach to the study of selective or attentional processes in perception is described in a paper by Kinchla (1980) entitled "The Measurement of Attention." The basic idea is that more efficient utilization of information from one source is often accompanied by a reduced ability to utilize information from another source. This sort of information tradeoff has been referred to as an "attention operating characteristic" (AOC) by Kinchla (Note 1)¹ and Sperling and Melchner (1978) and as a "performance operating characteristic" (POC) by Norman and Bobrow (1975). The form of such an information tradeoff (AOC or POC) is revealed by progressively encouraging an observer to "pay more attention to" one source of information than the other.

This approach was employed to examine how subjects utilized information from the two structural levels (global and local) of stimuli like that in Figure 1. The question was whether factors which increased the speed of utilization from one level would slow it from the other; that is, was there evidence of an information tradeoff (AOC or POC) consistent with some sort of selective or attentional process?

An Experimental Paradigm

The basic experimental paradigm consists of a series of trials of the sort illustrated in Figure 2. Each trial begins with the presentation of two target letters, one on the right of the display and one on the left. Next, a test pattern is presented defining two test letters, a large global letter, and many instances of a

FORCED-CHOICE PARADIGM

TARGET DEFINED → TEST PATTERN → RESPOND LEFT OR RIGHT

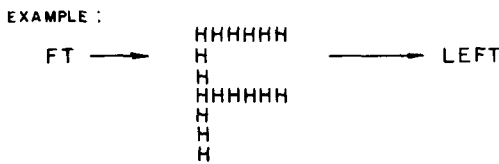


Figure 2. The forced-choice paradigm employed in the two experiments reported here. In this example, the *leftmost* target letter (F) is found at the global level of the test pattern, indicating a *left-hand* response.

smaller local letter. The observer's task is to decide which of the two target letters, left or right, matches one of the two test letters, and to indicate this by pressing a button on his left or right, respectively. For example, a left-button response is indicated in Figure 2, since the leftmost target letter (F) is defined at the global level of the test pattern. The two target letters are randomly selected on each trial and only one of them matches one of the two test letters. Thus, the appropriate response (left or right) is always indicated at one of the two structural levels of the test pattern (global or local), with an irrelevant letter at the other level. Test patterns which contain response information at the global level will be termed " S_G patterns," and those that contain it at the local level " S_L patterns." The type of test pattern presented on each trial was determined randomly: An S_G pattern was presented with probability $P(S_G)$ or an S_L pattern with probability $1 - P(S_G)$. These probabilities were fixed throughout a testing session and known to the observer.

Observers were told to "respond as rapidly as possible without making errors." Thus, the dependent measures of principal interest were the average response latency following the onset of an S_G pattern, denoted $L(S_G)$, or that following an S_L pattern, denoted $L(S_L)$. In particular, we were interested in whether progressively encouraging an observer to "attend" more to one level of structure or the other revealed a negative relation between $L(S_L)$ and $L(S_G)$.

A Binary Mixture Model

The following theoretical view was suggested both by the informal comments of our observers and our own phenomenal impressions of the task. It seemed that one had two alternative ways of "attending" to the display: one which facilitated perception of the global form and another which facilitated perception of the local form. In each case the cost of "seeing" one level of form rapidly and clearly seemed to be a slower or less vivid perception of the other level. These subjective impressions are represented more formally in the following Binary Mixture Model. The model is similar in general form to ones proposed earlier to characterize other perceptual phenomena (e.g., Falmange & Theios, 1969; Sperling & Melchner,

1978; Kinchla, Note 1). It has also recently come to our attention that Ward (in press) has also suggested that a binary mixture model may be appropriate for tasks of this sort, although he did not formally develop such a model or evaluate AOC functions.

Basic assumptions. The basic assumptions of the model can be summarized in two statements:

(1) At the start of each trial, an observer adopts one of *two alternative processing strategies*: a G strategy, which is optimal for processing global forms, or an L strategy, which is optimal for processing local forms.

(2) The mean response latency to an S_G stimulus pattern is t_G given a G strategy, or t'_G given an L strategy, where $t_G \leq t'_G$. Similarly, the mean latency to an S_L pattern is t_L given an L strategy, or t'_L given a G strategy, where $t_L \leq t'_L$.

Properties of the model. The preceding assumptions, which are graphically summarized at the top of Figure 3, lead directly to theoretical expressions for $L(S_G)$ and $L(S_L)$:

$$L(S_G) = P(G)t_G + P(L)t'_G \quad (1)$$

and

$$L(S_L) = P(L)t_L + P(G)t'_L \quad (2)$$

where $P(G)$ and $P(L)$ denote the probabilities of adopting each of the processing strategies, so that

Binary Mixture Model

		Stimulus Type	
		S_G	S_L
Processing Strategy	G	t_G	t'_L
	L	t'_G	t_L

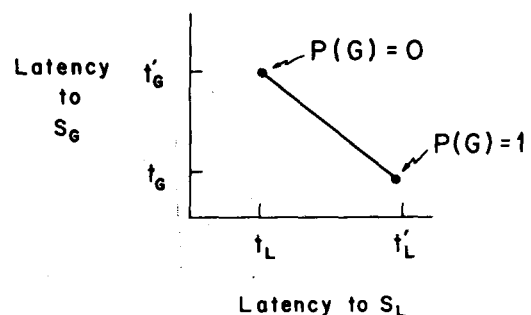


Figure 3. The binary mixture model depicts the average latency to each type of stimulus pattern as a mixture of trials, some when the processing strategy was optimal for that pattern and others when it was not. Shifts in the subject's tendency to utilize a global strategy, $P(G)$, generate the linear tradeoff (AOC) relating $L(S_G)$ and $L(S_L)$.

$P(G) = 1 - P(L)$. As an observer varies his tendency to adopt a G strategy ("attend to the global form"), his average response latencies to S_G and S_L patterns vary, respectively, from t_G and t_L' when $P(G) = 1$ to t_G' and t_L when $P(G) = 0$. This traces out the linear AOC relating $L(S_G)$ and $L(S_L)$ illustrated in the lower part of Figure 3. Each point along this AOC function corresponds to a particular *mixture* of trials, some when the processing strategy is optimal for the test pattern presented and others when it is not. The specific mixture for each stimulus type depends on $P(G)$, as indicated in Equations 1 and 2.

TWO EXPERIMENTS

Each of the following two experiments employed the experimental paradigm described earlier (Figure 2). In each experiment, observers were encouraged to vary their tendency to "attend" to one level of stimulus structure or the other. This was done in the first experiment by varying the proportion of trials on which the response information was at the global (rather than local) level, while in the second experiment the observers were simply told it was more important to attend to one level or the other.

Experiment 1

Method

Three paid observers were run under six experimental conditions defined by $P(S_G)$ equal to 0, .2, .4, .6, .8, or 1.

Each observer practiced the experimental task for about 1,500 trials, and then was tested during 10 daily 600-trial sessions. Before each successive 100-trial block the observer was told the proportion of S_G patterns in that block. All six values of $P(S_G)$ were employed in a randomly determined order each day, for a total of 1,000 trials in each condition. Stimuli were presented on a fast-decay CRT display under computer control. Each small letter on the display was composed from a conventional 6×8 matrix of illuminated dots and had an average luminance of about .85 cd/m². The display was viewed binocularly in the dark by a dark-adapted observer from a distance of about 80 cm, at which each 8×10 dot matrix subtended about .115 deg visual angle in height. The test patterns were composed from corresponding 6×8 arrays of the small letters; so the height of each large ("global") letter was about 1.5 deg visual angle, and its width about 1 deg.

At the beginning of each trial, two target letters were selected randomly from the set T, L, H, N, Z, Y, and X. These were displayed on the CRT about 10 deg apart until the observer felt ready

to proceed. When he pressed a button, the target letters were replaced by a fixation point in the center of the display, followed 500 msec later by a 100-msec-duration test pattern centered on the fixation point. The observer then had up to 2 sec to indicate his response. (If no response occurred within 2 sec, the trial was aborted.) Immediately following each response, a plus or minus appeared on the screen to indicate, respectively, a correct or incorrect response. The observer could then initiate the next trial by pressing a button.

Results

Table 1 presents average response latencies for the various combinations of stimulus type and $P(S_G)$, for each observer, along with the average of these values for all three observers. The most obvious feature of these data are the large shifts in latencies as $P(S_G)$ varied, with a negative relation between $L(S_G)$ and $L(S_L)$. This was confirmed statistically by a two-way analysis of variance on the data from conditions yielding estimates of both $L(S_G)$ and $L(S_L)$; specifically, stimulus type, S_G or S_L , was one factor and $P(S_G)$, equal to .2, .4, .6, or .8, the other, in a 2×4 analysis of variance on each observer's data. There was no significant main effect of stimulus type [$F(1,3) < 1$ for each observer], but there was a main effect of $P(G)$ [$F(3,3992) = 21.4, 39.9, \text{ and } 48.9$ for Observers 1, 2, and 3, respectively, with $p < .001$ in each case]. There was also a significant interaction between stimulus type and $P(G)$ [$F(3,3992) = 1,447, 522, \text{ and } 1,318$ for Observers 1, 2, and 3, respectively, with $p < .001$ in every case].

Error rates were very low, less than .03 in all conditions, with no discernible systematic difference between conditions.

Theoretical Analysis

The mixture model provides a theoretical framework for further evaluation of the data. In Figure 4, both the individual and average observer data from Table 1 are presented as AOC functions of the sort shown earlier in Figure 3. (Since the theoretical AOC for each observer is linear, the average observer data should also be described by a linear AOC function.) The four open points on each graph correspond to the $P(S_G) = .2, .4, .6, \text{ and } .8$ data used in our analyses of variance. Linear functions were fitted to these

Table 1
Experiment 1: Average Response Latencies, $L(S_G)$ and $L(S_L)$, in Milliseconds,
in Each Condition, for Each Observer, and Overall Averages

$P(S_G)$	Observer 1		Observer 2		Observer 3		Average	
	S_G	S_L	S_G	S_L	S_G	S_L	S_G	S_L
0		443		489		479		470
.2	740	510	738	554	798	547	759	537
.4	746	562	668	668	677	593	697	608
.6	548	708	651	733	599	682	599	708
.8	480	759	584	738	551	797	538	765
1.0	391		546		494		477	

(open) points on each graph by a least squares criterion. In each case, the four points are ordered along this line from $P(S_G) = .2$ in the upper left to $P(S_G) = .8$ in the lower right. This is consistent with each observer's adopting a G strategy ("attending to the global level") progressively more often as $P(S_G)$ increased from .2 to .8. The solid points on each graph are based on the $P(S_G) = 0$ and 1 data, since only estimates of, respectively, $L(S_L)$ and $L(S_G)$ could be obtained in those conditions. These (solid) points were used to estimate the end points of each theoretical AOC function under the assumption that $P(G) = 1$ when $P(S_G) = 1$ and $P(G) = 0$ when $P(S_G) = 0$. For example, if an observer knew that S_G would be presented on every trial, $P(S_G) = 1$, it seems reasonable to assume that he always "attended to the global level," $P(G) = 1$, and his mean latency to S_G (the solid point on the ordinate of each graph) can be taken as an estimate of t_G . A similar argument suggests that the solid point on the abscissa is an estimate of t_L . The intersection of these values and each AOC indicate the end points of the function and the values of t'_G and t'_L .

Estimates of t_G , t_L , t'_G , and t'_L corresponding to the four AOC functions in Figure 4 are presented in Table 2, along with the differences between t'_G and t_G and between t_L and t'_L . These differences can be interpreted as the average additional time required to utilize information from a specific level when an in-

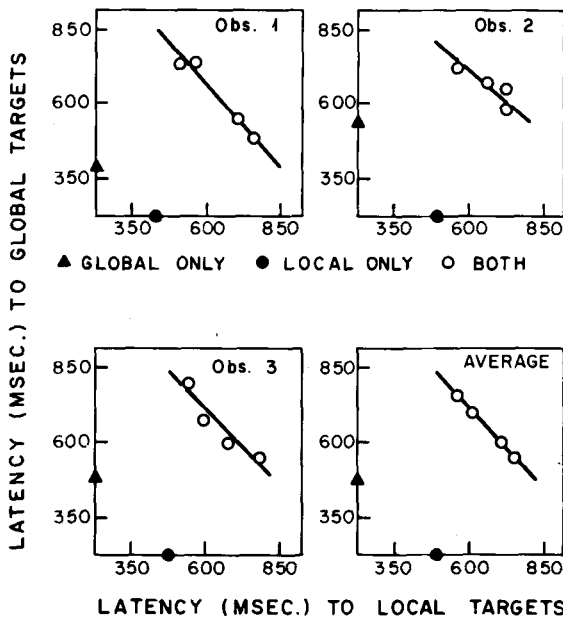


Figure 4. Individual and average data from Experiment 1. The solid points are from the $P(S_G) = 0$ (circle) and 1 (triangle) conditions in which only latencies to only one type of stimulus were measured. The open points from upper left to lower right, respectively, are from the $P(S_G) = .2, .4, .6,$ and $.8$ conditions in which both $L(S_G)$ and $L(S_L)$ were measured. The solid lines are theoretical AOC functions of the sort illustrated in Figure 3.

Table 2
Experiment 1: Estimates of Theoretical Mean Latencies and the Differences in Speed of Utilization When the Processing Strategy Matched (Unprimed) or Mismatched (Primed) the Stimulus Type (Indicated by the Subscript)

	Observer 1	Observer 2	Observer 3	Average
t_G	391	546	494	477
t'_G	850	822	846	839
$t'_G - t_G$	459	276	352	362
t_L	443	489	479	470
t'_L	850	803	824	826
$t'_L - t_G$	407	314	345	356

Table 3
Experiment 1: Estimates of $P(G)$ for $P(S_G) = .2, .4, .6,$ and $.8$

$P(S_G)$	Observer 1	Observer 2	Observer 3	Average
.2	.21	.26	.16	.24
.4	.25	.55	.41	.38
.6	.66	.69	.64	.66
.8	.79	.82	.87	.80

Note— $P(G)$ was assumed to equal 0 when $P(S_G) = 0$ and 1 when $P(S_G) = 1$.

appropriate processing strategy is employed. For example, the value of $t'_G - t_G$ for an average observer is 362 msec. This means that using a nonoptimal L strategy ("attending to the local level") when an S_G pattern is presented increases the average latency of a response by 362 msec over what it would be when using the optimal G strategy.

As an observer becomes more confident of finding the relevant stimulus information at a particular structural level, he is more likely to adopt the optimal processing strategy for that level. Table 3 presents numerical estimates of $P(G)$ for the data represented by open points in Figure 4. These estimates were obtained by first rewriting Equation 1 in the form

$$P(G) = \frac{t'_G - L(S_G)}{t'_G - t_G} \tag{3}$$

and substituting the appropriate estimates of t'_G and t_G from Table 2, along with an estimate of $L(S_G)$ corresponding to the point on each theoretical AOC closest to the observed data point for the particular $P(S_G)$ condition. Solving for $P(G)$ yielded the values in Table 3. In general, the observers seemed to "match" $P(G)$ to $P(S_G)$.

Experiment 2

Method

Three paid observers were run under five experimental conditions. Two of these conditions corresponded to the $P(S_G) = 0$ and 1 conditions in Experiment 1. $P(S_G) = .5$ in each of the other three conditions, which differed only in the instructions given to the observer: In a "mostly local" condition, the observer was told that "fast processing of local forms is most important"; in a "mostly global" condition, he was told that "fast processing of global

Table 4
Experiment 2: Average Response Latencies, $L(S_G)$ and $L(S_L)$, in Milliseconds, in Each Condition, for Each Observer, and Overall Averages

Instructions	Observer 1		Observer 2		Observer 3		Average	
	S_G	S_L	S_G	S_L	S_G	S_L	S_G	S_L
Only L		370		371		446		396
Mostly L	761	472	738	460	790	520	763	484
Both	539	543	604	612	649	640	597	598
Mostly G	452	731	452	772	512	775	472	759
Only G	350		351		424		375	

forms is most important"; and, in what will be termed the "both" condition, he was told that "fast processing of both levels of form is equally important."

The procedure was essentially the same as in Experiment 1. After about 1,500 practice trials, each observer was tested on 10 daily 500-trial sessions. Each of the five experimental conditions was employed in a randomly determined order during five successive 100-trial blocks each day, with the observers informed of the value of $P(S_G)$ in each block. This yielded a total of 1,000 trials under each condition.

Results

Table 4 presents average response latencies for the various combinations of stimulus type and experimental condition for each observer, along with the average of these values for all three observers. Again, as in Experiment 1, the most obvious feature of these data are the large shifts in latencies across the five experimental conditions and the negative relation between $L(S_L)$ and $L(S_G)$. This was confirmed statistically by a two-way (2×4) analysis of variance on the data from the three $P(S_G) = .5$ conditions; specifically, one factor was stimulus type (S_G or S_L) and the other factor, instructions ("mostly local," "mostly global," or "both"). There was no significant main effect of stimulus type [$F(1,2) < 1$ for each observer], but there was a main effect of instruction [$F(2,2994) = 175, 5.2, \text{ and } 5.6$ for Observers 1, 2, and 3, respectively, with $p < .01$ in each case]. There was also a significant interaction between stimulus type and instruction for each observer [$F(2,2994) = 2,395, 2,625, \text{ and } 2,462$, for Observers 1, 2, and 3, respectively, with $p < .001$ in each case].

Again, as in Experiment 1, error rates were very low, less than .03, in all conditions with no discernible systematic differences between conditions.

Theoretical Analysis

Figure 5 presents the data from Table 4 in the form of theoretical AOC functions similar to those obtained in Experiment 1 (Figure 4). The solid points on the axis correspond to the $P(S_G) = 0$ and 1 conditions, and the open points, to the $P(S_G) = .5$ conditions. Again, a linear function was fit to each set of open points using a least squares criterion, and in each case the points are ordered along this line from the "mostly local" condition in the upper left to the "mostly global" condition in the lower right.

Estimates of $t_G, t_L, t'_G,$ and t'_L (obtained as in Experiment 1) are presented in Table 5. It is clear that the data reflect linear trading relations quite comparable to those in Experiment 1. Table 6 presents estimates of $P(G)$ for the three instructional conditions (calculated as in Experiment 1). These imply that the instructions had a substantial effect on $P(G)$ of the same sort as $P(S_G)$ in Experiment 1; it is as if the more emphasis there was on "rapid processing of the global form," the greater the value of $P(G)$.

DISCUSSION AND CONCLUSIONS

The results of both experiments imply that one can "attend" to a particular level of structure in a visual image, utilizing information from that level more

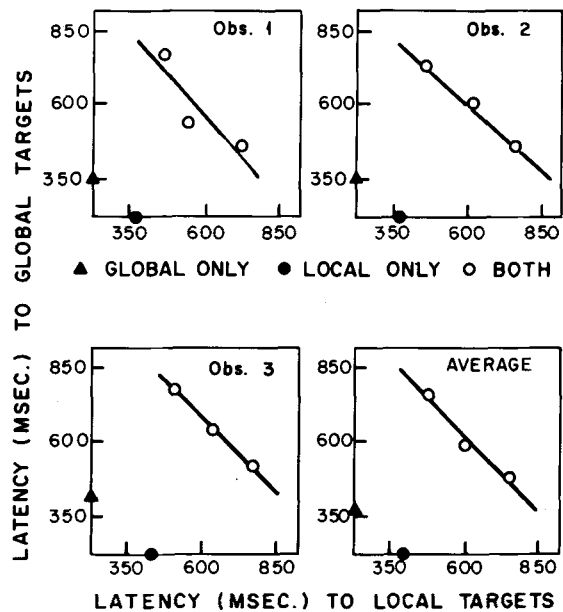


Figure 5. Individual and average data from Experiment 2. The solid points are from the $P(S_G) = 0$ (circle) and 1 (triangle) conditions in which latencies to only one type of stimulus were measured. The open points from upper left to lower right, respectively, are the "most local," "both," and "most global" conditions in which $P(S_G) = .5$ and both $L(S_G)$ and $L(S_L)$ were measured. The solid lines are theoretical AOC functions of the sort illustrated in Figure 3.

Table 5
Experiment 2: Estimates of Theoretical Mean Latencies and the Differences in Speed of Utilization When the Processing Strategy Matched (Unprimed) or Mismatched (Primed) the Stimulus Type (Indicated by the Subscript)

	Observer 1	Observer 2	Observer 3	Average
t_G	350	351	424	375
t'_G	818	813	829	845
$t'_G - t_G$	468	462	405	470
t_L	370	371	446	396
t'_L	789	883	861	846
$t'_L - t_L$	419	512	415	450

Table 6
Experiment 2: Estimates of $P(G)$ for Various Instruction Conditions

Instructions	Observer 1	Observer 2	Observer 3	Average
Mostly L	.16	.20	.10	.20
Both	.52	.50	.44	.52
Mostly G	.80	.78	.81	.76

Note— $P(G)$ was assumed to equal 0 and 1, respectively, in the $P(S_G)$ equal 0 ("L only") and $P(S_G)$ equal 1 ("G only") conditions.

rapidly at the cost of slower utilization of information from other levels. Furthermore, this tradeoff in speed is consistent with the linear AOC predicted by our simple mixture model. An observer can be induced to move his performance along this AOC either by shifts in $P(S_G)$ or by direct instructions to "attend more" to one level or another. Both manipulations could be interpreted as inducing a particular tendency $P(G)$, to adopt a global, rather than local, "processing strategy." Ward (1982) conducted a series of experiments using two-level stimulus patterns similar to that of Figure 1. His observers seemed better able to process information from one level if they had just processed information from the same level in a preceding pattern, which Ward termed a "level readiness effect." It could be argued that sequential effects of this sort contributed to the influence of $P(S_G)$ in our first experiment, since the larger the value of $P(S_G)$, the more often will a "preceding stimulus" be S_G . However, such sequential contingencies could not account for the instructional effects in our second experiment, since these contingencies were the same in all conditions in which $P(S_G) = .5$.

In defining the simple mixture model, we have been careful to use the more general concept of two alternative "processing strategies" (G or L), rather than specifying particular cognitive "mechanisms" or "processes" involved in such strategies. In Kinchla (1980), it was argued that instead of treating attention as a single entity, it is more useful to assume that a variety of cognitive mechanisms mediate selectivity in human information processing, any one of which could be considered "attentional." Thus, it

is the selectivity in processing evidenced by an empirical AOC function which underlies the concept of attention, rather than the particular theoretical mechanism mediating that selectivity.

The preceding point can be illustrated by briefly considering a few of the theoretical mechanisms which might be used to characterize our two "processing strategies." First of all, one might imagine the observer as selectively "tuning" or "sensitizing" his sensory system to different spatial frequency bands (e.g., see Davis & Graham, 1981, or Kramer, Yager, Graham, & Shaw, 1981). Faster processing of global information could occur if this "tuning" process favored low spatial frequency channels over higher ones. Here the earliest stages of encoding or representation selectively favor information about one level of form or the other. Alternatively, one might evoke a "feature extraction" process similar to that proposed by Rumelhart (1970) in which "features" are selectively "extracted" from a rapidly decaying iconic representation. Instead of the spatially defined "channels" in Rumelhart's representation, one could imagine orientation- and frequency-specific channels as being differentially important for processing global or local forms. Here selectivity occurs at a "feature extraction" stage following non-selective encoding or representation at the iconic level. Other models might represent selectivity as occurring when "features" are "organized" or "configured" (e.g., see Banks & Prinzmetal, 1976, or Pomerantz, 1981). Here "segmentation" or "partitioning" of features into subsets appropriate for the processing of local forms might be inconsistent with the processing of global ones. Finally, selectivity might be mediated by "ordered memory comparison processes" of the sort proposed by Falmagne and Theios (1969), Shiffrin and Geisler (1973), or Sternberg (1969). For example, our task (Figure 2) could be interpreted as a "two-probe" version of Sternberg's single-probe memory scanning task: our two test letters are probes to be compared with a memory set consisting of the two target letters. A strictly serial version of the comparison process could have the observer serially compare one of the probes (global or local) with the memory set (target letters), responding if a match occurred or initiating a second comparison of the other probe if it did not. Our two processing strategies (G or L) would then correspond to the probe used in the initial comparison. If it were appropriate for the stimulus, responses would occur faster than when a second comparison process was required. Again, the point here is simply that any one, or combination, of the preceding theoretical mechanisms could explain the AOC functions we have discovered.

It should be emphasized that our level of theoretical description does not distinguish between "processing strategies" involving a rapid "switching" be-

tween global and local levels of a test pattern and simultaneous "sharing" of some limited processing capacity. For example, the global processing strategy could be represented by a "switching" process in which most of the time was spent processing the global form or a "sharing" process whereby both levels were processed simultaneously but with more processing capacity allocated to the global level. Our mixture model would be appropriate in either case so long as each processing strategy (G or L) led to different mean latencies given each type of stimulus (S_G or S_L).

While it would be premature to argue for a specific selective mechanism, the idea of ordered memory comparison processes is quite appealing. First of all, the type of "interference effects" described earlier seem to be consistent with rapid encoding of *both* global and local forms at the letter-code level, since both may produce interference when the observer is told to ignore them (Hoffman, 1980; Miller, 1982) even at the high-speed end of the speed-accuracy tradeoff function (Boer & Keuss, 1982). This is why the authors of the last three cited papers argued that the level precedence effects were a "postperceptual" phenomenon. That is, it reflected what the observer "did" with the information rather than the order in which it was initially encoded (Miller, 1981b). Early encoding of both target letters is perfectly compatible with selectivity during memory comparisons, since the AOC would not reflect the relative speed at which the two test letters were encoded but, rather, the order in which they were compared with the test letters.

Furthermore, such ordered comparisons could in principle be based on any discernible aspect of multiple probe stimuli. For example, an observer might first compare the larger of two probes or the probe of a particular color, etc. While this would *not* be inconsistent with an ordering based on structural level, it does suggest that such selectivity may be only one instance of a more general aspect of human information processing.

Selective memory comparisons of this sort may underlie our phenomenal impression of "seeing" one level of structure faster or more vividly than another, not only in this specific task but much more generally. For example, suppose that just before the image of a clockface was flashed on a screen you had been asked to compare it with other types of clocks (wrist-watches, alarm clocks, etc.) or, alternatively, to describe the particular type of numerals on the clockface. Your phenomenal impression in each case should be quite different, primarily because you would be comparing the visual image with two different sets of items ("knowledge structures") in memory ("clocks" or "numerals"). There seems little doubt that humans can selectively retrieve information from longer term memory, and it may

be this sort of selection which determines what forms we "see" most vividly in an image.

Finally, it should be emphasized that the present experiments were *not* designed to evaluate the nature of errors on tasks of this sort. Rather, our instructions to the observers were designed to ensure very low error rates, so low that the effect of our experimental manipulations on error rates cannot really be assessed. Additional experiments are now being conducted in which "deadline" procedures have been employed to increase the speed pressure on an observer in order to produce higher error rates. In this way, it should be possible to develop a broader theoretical characterization that would include a treatment of errors as well as mean latencies. For example, it may be that only information from the "attended" level of structure can be effectively utilized when the observer is under sufficient speed pressure. Thus, factors which seem to influence $P(G)$ in the experiments reported in this paper would also influence error patterns in the proposed experiments. Our use of a deadline procedure is similar to that in the earlier cited paper by Boer and Keuss (1982), although they used a different perceptual task and were interested solely in speed-accuracy tradeoff functions. We are using the same task as in the present paper except for a variable deadline. This will allow us to increase speed pressure to produce higher error rates. If the general idea behind the binary mixture model is appropriate, it should be possible to generate linear *accuracy* AOC functions similar to the *latency* AOC functions of the present study. This conclusion follows directly if we simply add a *third assumption to the two previously stated basic assumptions of the model*:

(3) The probability of an incorrect response to an S_G stimulus pattern is e_G , given a G strategy, or e'_G , given an L strategy, where $e_G \leq e'_G$. Similarly, the probability of an erroneous response to an S_L pattern is e_L , given an L strategy, or e'_L , given a G strategy, where $e_L \leq e'_L$.

In other words, an observer not only utilizes information more rapidly if his processing strategy (G or L) matches the stimulus type (S_G or S_L), but he will also be more accurate. We can now write the following expressions for the conditional probability of an error, denoted by E, given S_G or S_L :

$$P(E | S_G) = P(G)e_G + P(L)e'_G \quad (3)$$

$$P(E | S_L) = P(L)e_L + P(G)e'_L \quad (4)$$

This implies a negative linear relation between $P(E | S_G)$ and $P(E | S_L)$ of the same form as that between $L(S_G)$ and $L(S_L)$ shown in Figure 3, only here variation in $P(G)$ moves performance from the point e_G, e'_L when $P(G)$ equals 1 to the point e'_G, e_L when $P(G)$ equals 0. Thus, if the simple mixture model is appropriate, it should

be possible to recover consistent estimates of $P(G)$ from both latency and error data.

REFERENCE NOTE

1. Kinchla, R. *An attention operating characteristic in vision* (Tech. Rep. 29). Hamilton, Ontario, Department of Psychology, McMaster University, 1969.

REFERENCES

- BANKS, W., & PRINZMETAL, W. Configurational effects in visual information processing. *Perception & Psychophysics*, 1976, **19**, 361-367.
- BOER, L., & KEUSS, P. Global precedence as a postperceptual effect: An analysis of speed-accuracy tradeoff functions. *Perception & Psychophysics*, 1981, **31**, 358-366.
- DAVIS, E., & GRAHAM, N. Spatial frequency uncertainty effects in the detection of sinusoidal gratings. *Vision Research*, 1981, **21**, 705-712.
- FALMANGE, J., & THEIOS, J. On attention and memory in reaction time experiments. In W. G. Koster (Ed.), *Attention and Performance II*. Amsterdam: North Holland, 1969.
- FU, K. *Syntactic methods in pattern recognition*. New York: Academic Press, 1974.
- HOFFMAN, J. Interactions between global and local levels of form. *Journal of Experimental Psychology: Human Perception and Performance*, 1980, **6**, 222-234.
- KINCHLA, R. A. Detecting target elements in multi-element arrays: A confusability model. *Perception & Psychophysics*, 1974, **15**, 149-158.
- KINCHLA, R. The role of structural redundancy in the perception of visual targets. *Perception & Psychophysics*, 1977, **22**, 19-30.
- KINCHLA, R. The measurement of attention. In R. S. Nickerson (Ed.), *Attention and performance VIII*. Hillsdale, N.J.: Erlbaum, 1980.
- KINCHLA, R. A., & WOLFE, J. The order of visual processing: "Top down," "bottom up," or "middle out." *Perception & Psychophysics*, 1979, **25**, 225-231.
- KRAMER, P., YAGER, D., GRAHAM, N., & SHAW, M. Recognition of spatial frequency: Effect of set size. *Supplement to Investigative Ophthalmology and Visual Science*, 1981, **20**, 124.
- MARTIN, M. Local and global processing: The role of sparsity. *Memory & Cognition*, 1979, **7**, 476-484.
- MILLER, J. Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, **7**, 1161-1174. (a)
- MILLER, J. Global precedence: Information availability or use? Reply to Navon. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, **7**, 1183-1185. (b)
- MILLER, J. Global precedence in attention and decision. *Journal of Experimental Psychology*, 1982, **18**, 240-275.
- NAVON, D. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 1977, **9**, 353-383.
- NORMAN, D., & BOBROW, D. On data-limited and resource-limited processes. *Cognitive Psychology*, 1975, **7**, 44-64.
- PALMER, S. Visual perception and world knowledge. In D. A. Norman & D. E. Rumelhart (Eds.), *Explorations in cognition*. San Francisco: Freeman, 1975.
- POMERANTZ, J. Perceptual organization in information processing. In M. Kubovy (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981.
- POMERANTZ, J., & SAGER, L. Asymmetric integrality with dimensions of visual pattern. *Perception & Psychophysics*, 1975, **18**, 460-466.
- RUMELHART, D. A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology*, 1970, **7**, 191-218.
- SHIFFRIN, R., & GEISLER, W. Visual recognition in a theory of information processing. In R. Solso (Ed.), *Contemporary issues in cognitive psychology: The Loyola Symposium*. Washington, D.C.: Winston, 1973.
- SPELTING, G., & MELCHNER, M. The attention operating characteristic: Examples from visual search. *Science*, 1978, **202**, 315-318.
- STERNBERG, S. Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist*, 1969, **57**, 421-457.
- SWETS, J. A., & KRISTOFFERSON, A. B. Attention. In P. H. Mussen & M. R. Rosenzweig (Eds.), *Annual review of psychology*, Palo Alto, Calif: Annual Reviews, 1970.
- WARD, L. M. Determinants of attention to local and global features of visual form. *Journal of Experimental Psychology: Human Perception and Performance*, 1982, **8**, 526-581.

NOTE

1. The paper by Kinchla (Note 1) which introduced the idea of an attention operating characteristic is described in Swets and Kristofferson (1970).

(Manuscript received June 23, 1982;
revision accepted for publication October 19, 1982)